Computational Geometry 46 (2013) 494–519

Contents lists available at SciVerse ScienceDirect

Computational Geometry: Theory and Applications

www.elsevier.com/locate/comgeo

A modified Frank–Wolfe algorithm for computing minimum-area enclosing ellipsoidal cylinders: Theory and algorithms

S. Damla Ahipaşaoğlu [a,*,1], Michael J. Todd [b,1]

[a] Department of Management, Management Science Group, London School of Economics, Houghton Street, London WC2A 2AE, UK
[b] Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853, USA

ARTICLE INFO

Article history:
Received 30 June 2010
Accepted 10 November 2011
Available online 16 November 2011

Keywords:
Linear convergence
Frank–Wolfe algorithm
Minimum-volume ellipsoids
Minimum-area cylinders
D-optimality
D_k-optimality

ABSTRACT

We study a first-order method to find the minimum cross-sectional area ellipsoidal cylinder containing a finite set of points. This problem arises in optimal design in statistics when one is interested in a subset of the parameters. We provide convex formulations of this problem and its dual, and analyze a method based on the Frank–Wolfe algorithm for their solution. Under suitable conditions on the behavior of the method, we establish global and local convergence properties. However, difficulties may arise when a certain submatrix loses rank, and we describe a technique for dealing with this situation.

1. Introduction

We study the problem of finding an ellipsoidal cylinder containing a finite set of points in $\mathbb{R}^n$, such that its cross-section with a fixed k-dimensional subspace has minimum area. This is a generalization of the minimum-volume enclosing ellipsoid (MVEE) problem, which has been much studied, with applications in data analysis and (via its dual) the D-optimal design problem in statistics.

The minimum-area enclosing ellipsoidal cylinder (MAEC) problem has also been widely studied, mainly because its dual is another optimal design problem in statistics, where now one is interested in estimating just k out of n parameters in a regression problem by choosing the design points optimally in some sense. See Fedorov [8], Silvey and Titterington [17], Atwood [2,3], and Pukelsheim [15] for more details.

Another application of the problem arises when studying possible collisions of bodies in $\mathbb{R}^3$. Suppose for a given body we are given a large sample of time–space readings $(t_i; x_i)$. Then by computing a minimum-area enclosing ellipsoidal cylinder for these points ($n = 4, k = 3$), we obtain a conservative model for the body of an ellipsoid moving in a uniform direction at uniform speed. If we obtain such models for two bodies, a sufficient condition for them not colliding is that the corresponding ellipsoidal cylinders do not intersect, and this can be checked after a transformation of variables by solving a single convex quadratic trust-region subproblem in $\mathbb{R}^3$, which can be done efficiently [6,18]. If the bodies are moving under the influence of a gravitational field, a tighter model can be achieved by approximating the movement of the bodies by parabolas, which can easily be performed by replacing x by $x - (1/2)gt^2$, where $g \in \mathbb{R}^3$ is a vector corresponding to the gravitational field, and then proceeding as above.

* Corresponding author.
E-mail address: ahipasa@gmail.com (S.D. Ahipaşaoğlu).
[1] Supported in part by NSF through grant DMS-0513337 and ONR through grant N00014-08-1-0036.

0925-7721/$ – see front matter
doi:10.1016/j.comgeo.2011.11.004

Our interest in the MAEC problem is mainly algorithmic: we study a first-order method based on the Frank–Wolfe method [9] with Wolfe's away steps [20], which was introduced in the context of optimal design by Atwood [3]. The Frank–Wolfe method is also known as the conditional gradient method — see Polyak [14] and Bertsekas [4]. However, no detailed analysis of this method has been performed, and in certain cases it breaks down unless modified to keep an appropriate matrix positive definite. The reason for our interest in this algorithm is that each iteration is very cheap and the memory requirements are minimal. By contrast, interior-point methods may be more efficient for small-scale problems, but quickly run into memory difficulties with large-scale problems, as seen in the MVEE case in [1]. Even general first-order methods can be expensive, as discussed in Section 3.1 below.

We show that under reasonable conditions on the iterates produced by the algorithm, global complexity estimates and local convergence properties can be established. These conditions require that a certain principal submatrix of a positive semidefinite matrix produced by the algorithm remain positive definite. Global convergence under suitable conditions (ensuring the positive definiteness of this submatrix) was established by Atwood [3], while complexity estimates of related methods (usually omitting the away steps and assuming a Lipschitz continuous gradient of the objective function) were given by Levitin and Polyak [13] and Wolfe [20]; see also [14,4]. Linear convergence of the Frank–Wolfe method with away steps was proved by Guélat and Marcotte [10] under strong conditions, and for the MVEE problem (which does not satisfy these conditions) by Ahipaşaoğlu, Sun, and Todd [1], but the present problem is more general still.

We also provide a technique that allows the iterations to proceed when rank deficiency occurs; although we have no guarantee of convergence in this case, the method appears to work in practice. Finally, some computational results for large random problems are given.

The paper is organized as follows. In the next section, we state simple convex formulations of the MAEC problem and its dual. Although previous papers have included formulations with some convexity properties, they have not been fully convex. Section 3 describes the basic algorithm, which was introduced by Atwood [3] in 1973. We prove global and local convergence results in Section 4. The case of rank-deficiency of a critical submatrix is discussed in Section 5, and Section 6 contains the results of our computational study. We conclude in Section 7 with some final remarks. A short list of notation is included in Appendix A for quick reference. The details of some of the proofs are presented in Appendices B–D.

2. Problem formulation and duality

In this section we provide simple convex formulations of the MAEC problem and its dual. Previous formulations, such as that in Silvey and Titterington [17], have not been convex in all the variables (see problem (P') below), although they have some convexity properties, and the dual problem involved a Schur complement rather than our simpler formulation (D).

We also relate these formulations to earlier ones and prove duality results. The section ends by defining notions of approximate optimality, which will be used in our algorithms.

2.1. Problem definition, convex formulation

Suppose we are given a matrix $X = [x_1, x_2, \dots, x_m] \in \mathbb{R}^{n\times m}$ whose columns, the points $x_1, \dots, x_m$, span $\mathbb{R}^n$. Let $X = \left[\begin{smallmatrix} Z \\ Y \end{smallmatrix}\right]$ be a partition of X, where $Z \in \mathbb{R}^{(n-k)\times m}$ and $Y \in \mathbb{R}^{k\times m}$. If H' is a symmetric matrix of order k that is positive definite (we write $H' \succ 0$) and E is a matrix of order $k \times (n-k)$, then the set

$$\mathscr{C}\big(0; E; H'\big) := \big\{[z; y]:\ z \in \mathbb{R}^{n-k},\ y \in \mathbb{R}^k,\ (y + Ez)^T H'(y + Ez) \leqslant k\big\}$$

is a central (i.e., centered at the origin) ellipsoidal cylinder whose intersection with the subspace

$$\Pi := \big\{[z; y] \in \mathbb{R}^n:\ z \in \mathbb{R}^{n-k},\ y \in \mathbb{R}^k,\ z = 0\big\}$$

has volume $(\det H')^{-1/2}$ times that of a Euclidean ball in $\mathbb{R}^k$ of radius $\sqrt{k}$. H' determines the shape of the cross-section and E the "directions of the axes" of the cylinder. Observe that the cylinder can also be written as a "degenerate" ellipsoid: indeed

$$\mathscr{C}\big(0; E; H'\big) = \big\{x \in \mathbb{R}^n:\ x^T H x \leqslant k\big\}, \tag{2.1}$$

where H is positive semidefinite with rank k:

$$H := \begin{bmatrix} E^T H' E & E^T H' \\ H' E & H' \end{bmatrix}, \quad \text{using the relation } y + Ez = [E, I](z \quad y).$$

Hence, finding a central ellipsoidal cylinder, which contains the columns of X and has minimum-volume (area) intersection with Π, amounts to solving

$$\min_{H' \succ 0, E}\ \bar{f}(H', E) := -\ln\det H'$$

$$(y_i + E z_i)^T H'(y_i + E z_i) \leqslant k, \quad i = 1, \dots, m, \tag{P'}$$